Articles

Molecular Cloning and Characterization of Two Rat Renal Kallikrein Genes[†,‡]

Yin-Peng Chen,[§] Julie Chao,[∥] and Lee Chao*[,§]

Departments of Biochemistry and Pharmacology, Medical University of South Carolina, Charleston, South Carolina 29425

Received March 1, 1988; Revised Manuscript Received May 25, 1988

ABSTRACT: Kallikreins compose a multigene family coding for a subgroup of serine proteases, which are involved in the processing of bioactive peptides. Two rat kallikrein-related genes, RSKG-7 (rat submandibular gland kallikrein gene 7) and RSKG-3, have been cloned and their sequences analyzed. RSKG-7 is approximately 4200 bases in length and consists of five exons and four introns. The 5′ end region contains the variant CATAT box and TTTAAA box; the 3′ end region contains the polyadenylation signal AATAAA. This gene encodes a putative 28 935-dalton preproenzyme of 261 amino acids (aa). The active enzyme consists of 237 aa and is preceded by a deduced signal peptide of 18 aa and a profragment of 6 aa. RSKG-3 is highly homologous to RSKG-7 in terms of its sequence and structure; it encodes a 28 730-dalton prepropeptide consisting of a signal peptide of 18 aa, a profragment of 6 aa, and an active peptide of 235 aa. Sequence comparisons of RSKG-7, RSKG-3, and other kallikrein-related enzymes reveal the key amino acid residues needed for both serine protease activity (His/Asp/Ser) and kallikrein-like cleavage specificity at basic amino acids. Northern blot analyses using specific oligonucleotide probes demonstrate that, among the 12 tissues studied, RSKG-7 and RSKG-3 are expressed in the rat kidney and submandibular gland. Castration of male rats results in a decrease in submandibular gland RSKG-7 mRNA, which can be restored to the normal level by treatment with thyroxine or testosterone. On the other hand, neither castration nor hormonal manipulation affects RSKG-7 mRNA levels in the kidney.

Many proteins are synthesized as inactive precursors or zymogens that are subsequently converted into physiologically active forms by selective enzymatic cleavage of peptide bonds. This limited proteolysis is an important control element that can initiate new physiological functions or regulate preexisting ones. Typical examples are the processes of blood coagulation, fibrinolysis, complement activation, digestion, and hormone production (Neurath & Walsh, 1976). A great variety of enzymes including serine, zinc, thiol, and carboxyl proteases are produced to accomplish these tasks. Kallikreins are a subgroup of serine proteases that are involved in the processing of bioactive peptides by which diverse functions such as blood pressure regulation and maturation of growth factors and peptide hormones are accomplished (Schachter, 1980).

Kallikreins and kallikrein-like enzymes have been identified in and isolated from such diverse tissue as pancreas, salivary gland, gastrointestinal mucosa, sweat gland, pituitary gland, kidney, spleen, and brain (MacDonald & Ashley, 1985; Chao et al., 1983, 1984a, 1987; Mason et al., 1983). They are acidic glycoproteins with p*I*'s near 4 and molecular weights ranging from 25 000 to 43 000. They all have a charge relay system composed of histidine, aspartic acid, and serine residues, which form the crucial spatial configurations of the enzyme active site. Kallikreins are characterized by their ability to release vasoactive kinin peptides from kininogen substrates by limited proteolysis. However, the kininogenase activity of different kallikreins is highly variable. The most thoroughly characterized of these kallikrein-like proteins are epidermal growth factor binding protein (EGF-BP) and the γ subunit of nerve growth factor (γ-NGF), which are responsible for the processing and maturation of epidermal growth factor and nerve growth factor, respectively (Frey et al., 1979; Berger & Shooter, 1977).

We have identified three kallikrein-related proteins, tissue kallikrein, arginine esterase A, and tonin, in rat kidney (Chao, 1983; Chao et al., 1984b; Woodley-Miller et al., 1987). The two former enzymes release kinins, which are extremely potent stimuli to renal biochemical and physiological events (Margolius, 1984). Studies have shown that urinary secretion of tissue kallikrein is reduced in essential hypertension (Margolius et al., 1974), and it has been suggested that the kallikrein–kinin system in the kidney is involved in the etiology of this disease through its effects on sodium and water excretion (Levinsky, 1979). Tonin potentially plays an important role in cardiovascular function due to its ability to directly convert angiotensinogen to angiotensin II, a potent vasoconstrictive peptide (Boucher et al., 1974).

We have demonstrated that there are approximately 20 kallikrein-related genes in rat (Gerald et al., 1986a). These family members are highly homologous in sequences and structures, which strongly suggests that they evolved from a common ancestor gene by gene duplications. The present study describes the sequencing and characterization of two novel rat renal kallikrein genes, RSKG-7 and RSKG-3. The hormonal control of the expression of RSKG-7 is also studied.

Experimental Procedures

DNA Subcloning and Sequencing. Two clones, RSKG-7 and RSKG-3, of the Charon-4A-based rat submandibular gland genomic library (Gerald et al., 1986a) were isolated and subcloned into M13 phage. Single-stranded phage DNAs were prepared as described by Schreir and Cortese (1979) and were

[†] This work was supported by NIH Grants HL29397 and HL33552 and American Cancer Society Grant BC556.

[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02837.

* Address correspondence to this author.

[§] Department of Biochemistry.

[∥] Department of Pharmacology.

0006-2960/88/0427-7189$01.50/0

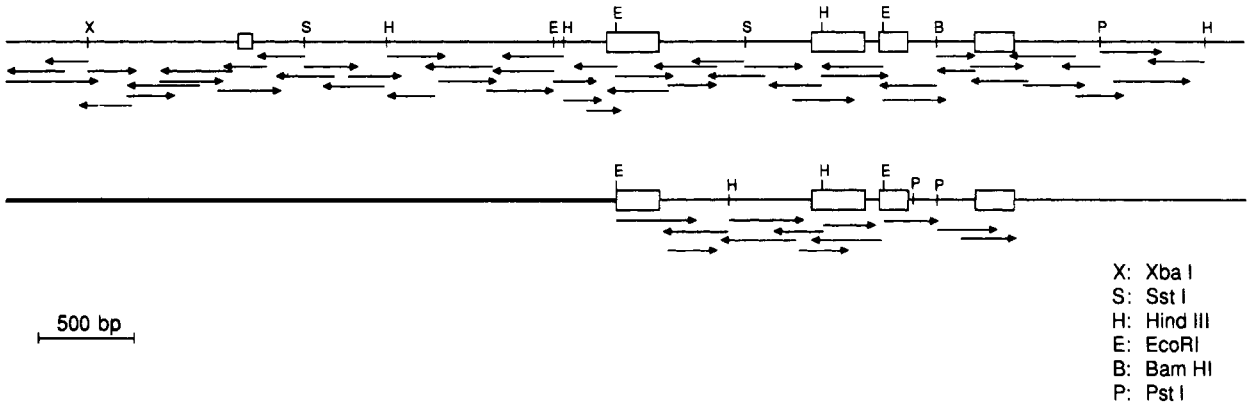


FIGURE 1: Sequencing strategy of rat kallikrein genes RSKG-7 and RSKG-3. The arrows denote the direction and length of sequences obtained from each sequencing reaction. The letters represent restriction enzyme sites utilized for gene mapping and subcloning. The open boxes show the exon regions, and the introns are indicated by thin lines. The cloning vector Charon-4A DNA is indicated by thick lines.

sequenced by the dideoxy chain termination method (Sanger et al., 1977). Primers for the sequencing reactions were either obtained commercially (BRL) or synthesized by using the ABI DNA synthesizer Model 380B.

RNA Preparations. Sprague-Dawley rats weighing 250 g were anesthetized and perfused with phosphate-buffered saline. The submandibular gland, adrenal gland, kidney, pancreas, prostate, lung, liver, thymus, pituitary gland, brain, testis, and heart were removed and homogenized in a solution of 4 M guanidine thiocyanate, 0.5% sodium lauryl sarcosine, 25 mM sodium citrate, and 0.1 M β-mercaptoethanol. Total RNAs were pelleted by centrifuging the tissue homogenates through a CsCl gradient (Davis et al., 1986). Submandibular gland polyadenylated RNA was isolated from the total RNA extract by two cycles of chromatography on a type III oligo(dT)–cellulose column (Collaborative Research).

Synthetic Oligodeoxyribonucleotide Probes. Oligodeoxyribonucleotides, each complementary to a specific region of RSKG-7, RSKG-3, or tissue kallikrein mRNA, were synthesized. The oligonucleotides were then end labeled (specific activity of approximately 10^8 cpm/μg) by using a 5′ terminus labeling system (BRL). Briefly, 100 pmol of oligonucleotide, 10 units of T4 polynucleotide kinase, 100 μCi of [γ-^{32}P]ATP, and 6 μL of 5× forward buffer were added to a total volume of 30 μL; this mixture was incubated at 37 °C for 1 h. The reaction was terminated by the addition of EDTA (to 12.5 mM). Free [γ-^{32}P]ATP was removed by a G-25 spun column.

Dideoxynucleotide Sequencing of RSKG-3 mRNA. A dideoxynucleotide RNA sequencing technique was developed by a modification of previously published techniques for cDNA synthesis (Sood et al., 1981) and the dideoxynucleotide sequencing of immunoglobulin RNA (Hamlyn et al., 1978). RSKG-3-specific ^{32}P-labeled oligonucleotide primer (0.5 μg in 1 μL) and 10 μg of salivary gland poly(A+) RNA in 10 μL of annealing buffer (50 mM Tris-HCl, pH 8.3; 50 mM NaCl; and 5 mM dithiothreitol) were heated to 60 °C for 3 min and then frozen immediately in a dry ice–ethanol bath. The mixture was thawed on ice to allow the annealing of the primer and the RSKG-3 mRNA. Each dideoxynucleotide sequencing mixture containing 2 μL of the primer–RNA template solution, 1 μL of 0.6 mM ddNTP, 1 μL of dNTP mix (2 mM of each dNTP), and 1 μL of 12.5 units of AMV reverse transcriptase in 50 mM Tris-HCl (pH 8.3), 50 mM NaCl, 5 mM dithiothreitol, and 24 mM $MgCl_2$ buffer was incubated at 48 °C for 15 min. The same amount of the reverse transcriptase was then added to each reaction mixture, and the mixtures were further incubated for another 15 min. The RNA sequence was resolved in a 8% polyacrylamide gel.

Hormone Treatment. Castrated Sprague-Dawley male rats and sham-operated controls were used in groups of six or seven animals. Animals were housed in groups of four in a light-controlled (12 h of light, 12 h of darkness), constant temperature (25 °C) room. All rats were fed with a normal lab chow (0.18 mequiv of Na^+/g, 0.2 mequiv of K^+/g; Wayne Pet Food Division, Chicago, IL) and allowed free access to tap water. Three weeks after surgery by the supplier, rats began to receive subcutaneous injections of 17α-methyl testosterone (40 mg/kg) or thyroxine (0.3 mg/kg) every other day for two weeks. At the end of this period total RNAs were isolated from the submandibular gland and kidney.

Northern Blot and Dot Blot Analysis. In the Northern blot analysis, total RNAs from submandibular gland, adrenal gland, kidney, pancreas, prostate, lung, liver, thymus, pituitary gland, brain, testis, and heart were resolved in a 1.5% agarose gel containing 0.66 M formaldehyde according to the method of Davis et al. (1986). The RNAs were transferred onto Zeta-probe membrane (Bio-Rad) by blotting the gel in 20× SSC (3 M NaCl, 0.3 M trisodium citrate) for 12–16 h. In the dot blot analysis, total RNAs were applied directly to the Zeta-probe membrane, following the procedures described by White et al. (1982). The membranes were baked at 80 °C in vacuo for 2 h, prehybridized in a solution containing 5× SSPE (0.75 M NaCl, 0.5 M NaH_2PO_4, 5 mM EDTA), 10× Denhardt's solution (0.2% of Ficoll 400, polyvinylpyrrolidone, bovine serum albumin), 0.5% sodium lauryl sarcosine, and 100 μg/mL denatured salmon sperm DNA at 55 °C for 4 h. Hybridization was performed at 50 °C overnight in the same solution containing 3 × 10^5 cpm/mL ^{32}P-labeled RSKG-7, RSKG-3, or tissue kallikrein specific oligonucleotide probe. The membranes were washed in three changes of 6× SSC and 0.5% sodium lauryl sarcosine at room temperature for 15 min and in two changes of the same solution at 42 °C for 15 min. Membranes were then exposed to X-ray film at −70 °C with an intensifying screen.

RESULTS

Primary Structures of RSKG-7 and RSKG-3. A series of restriction fragments from the two kallikrein-related gene clones RSKG-7 and RSKG-3 were subcloned into M13 phage. The overlapping subclones were constructed for the precise determination of the gene structure and as a source of DNA for sequencing. Both strands of RSKG-7 and the 5′ and 3′ flanking sequences, each approximately 1000 nucleotides in length, were completely sequenced (Figure 1). The variant TTTAAA box and CATAT box located at −21 and −71, respectively, are likely candidates of transcriptional regulatory

signals (Figure 2). Messenger RNA sequencing (data not shown) and the comparison of the nucleotide sequence of RSKG-7 with the known cDNA sequence of rat pancreatic kallikrein (Swift et al., 1982) and mouse kallikrein-related gene sequences (Evan & Richards, 1985) facilitated the accurate mapping of the junctions of introns and exons. RSKG-7 is composed of five exons and four introns. All of the introns begin with a GT dinucleotide and end with AG, sequences thought to be necessary for correct RNA splicing (Lerner et al., 1980). The gene contains approximately 4200 nucleotides, of which 873 bases are transcribed into mRNA. RSKG-7 mRNA sequence contains a consensus ribosome binding sequence (CCUGCUGC) in the 5′-noncoding region (42 bases) and a UGA stop codon and a AAUAAA polyadenylation signal 15 bases upstream of the polyadenylation site in the 3′-noncoding region (45 bases). The coding region consists of 783 nucleotides encoding a predicted prepropeptide of 261 amino acids with a M_r of 28 935.

We have also sequenced a truncated genomic clone, RSKG-3, from a conserved *Eco*RI site in the second exon of the gene to the polyadenylation site (Figures 1 and 2). The exon sequence upstream of the *Eco*RI site was determined by sequencing RSKG-3 mRNA using the dideoxynucleotide sequencing method, which was also used to map the transcriptional initiation sites of RSKG-3, tissue kallikrein gene (Figure 3), and RSKG-7 (data not shown). Submandibular gland mRNA was used for sequencing because, among the tissues studied, it expressed all three genes at the highest level. The priming sites of all three sequencing reactions were at the same region in the second exon. The autoradiograph (Figure 3) demonstrates that the RSKG-3 transcript is six nucleotides shorter than that of tissue kallikrein and that of RSKG-7 (data not shown) due to the deletion in the second exon, a fact which was also demonstrated by DNA sequencing (Figure 2). In the top portion of the sequence profiles of RSKG-3 (Figure 3), there are two clusters of strong termination bands across all four lanes (with the exception of G lanes, due to the high amount of dideoxy-GTP used in the reactions). The same profile also appears in tissue kallikrein and RSKG-7 mRNA dideoxy sequencing. It is unlikely that every strong termination would represent a transcriptional initiation site. The sequence profile has virtually revealed that AMV reverse transcriptase has a tendency to run off the template prematurely when it reaches the 5′ end of the mRNAs. Therefore, the results show that the last strong runoff in the autoradiograph (Figure 3) is the only transcriptional initiation site of the gene.

The relatively low abundance of RSKG-3 transcripts in the rat submandibular gland and the appearance of the secondary structures in the 5′ end have been major obstacles hindering our sequencing efforts. Thirty nucleotides in the 5′-noncoding region of the RSKG-3 transcript remain undetermined. However, the coding sequence and the gene organization of RSKG-3 have been revealed by a combination of mRNA and DNA sequence analyses. RSKG-3 is highly homologous to RSKG-7 in terms of the length and the organization of the introns and the exons (Figures 1 and 2). The coding region of RSKG-3 consists of 777 nucleotides encoding a prepropeptide of 259 aa with a M_r of 28 730.

RSKG-7 and RSKG-3 Encode Kallikrein-Related Serine Proteases. The prepropeptides encoded by RSKG-7 and RSKG-3 contain hydrophobic signal peptides of 18 amino acids and profragments of 6 amino acids (Figure 4). The short profragment is consistent with the isolation of an inactive kallikrein that can be proteolytically activated without significant change in molecular weight (Schachter, 1980). The amino acids histidine-41, aspartate-96, and serine-189 that constitute the charge relay system of serine proteases are conserved in RSKG-7 as well as in RSKG-3. The coding region of RSKG-7 and RSKG-3 share 85–87% sequence identity with that of rat tissue kallikrein. When amino acid sequences are compared (Figure 4), they share 76–88% sequence identity with the rat kallikrein-related proteins (PS, S2, and S3) and 64–70% identity with the mouse kallikrein-related proteins. These results clearly indicate that RSKG-7 and RSKG-3 are members of the kallikrein multigene family.

The residues that are known to line the substrate-binding pocket (Tschesche et al., 1979) are illustrated in Figure 4. RSKG-7, RSKG-3, PS kallikrein (Ashley & MacDonald, 1985), and other kallikrein-related proteins show a significant variation in these particular residues compared with the rest of the sequence. It is therefore tempting to speculate that RSKG-7 and RSKG-3 may have substrate specificities different from those of other family members.

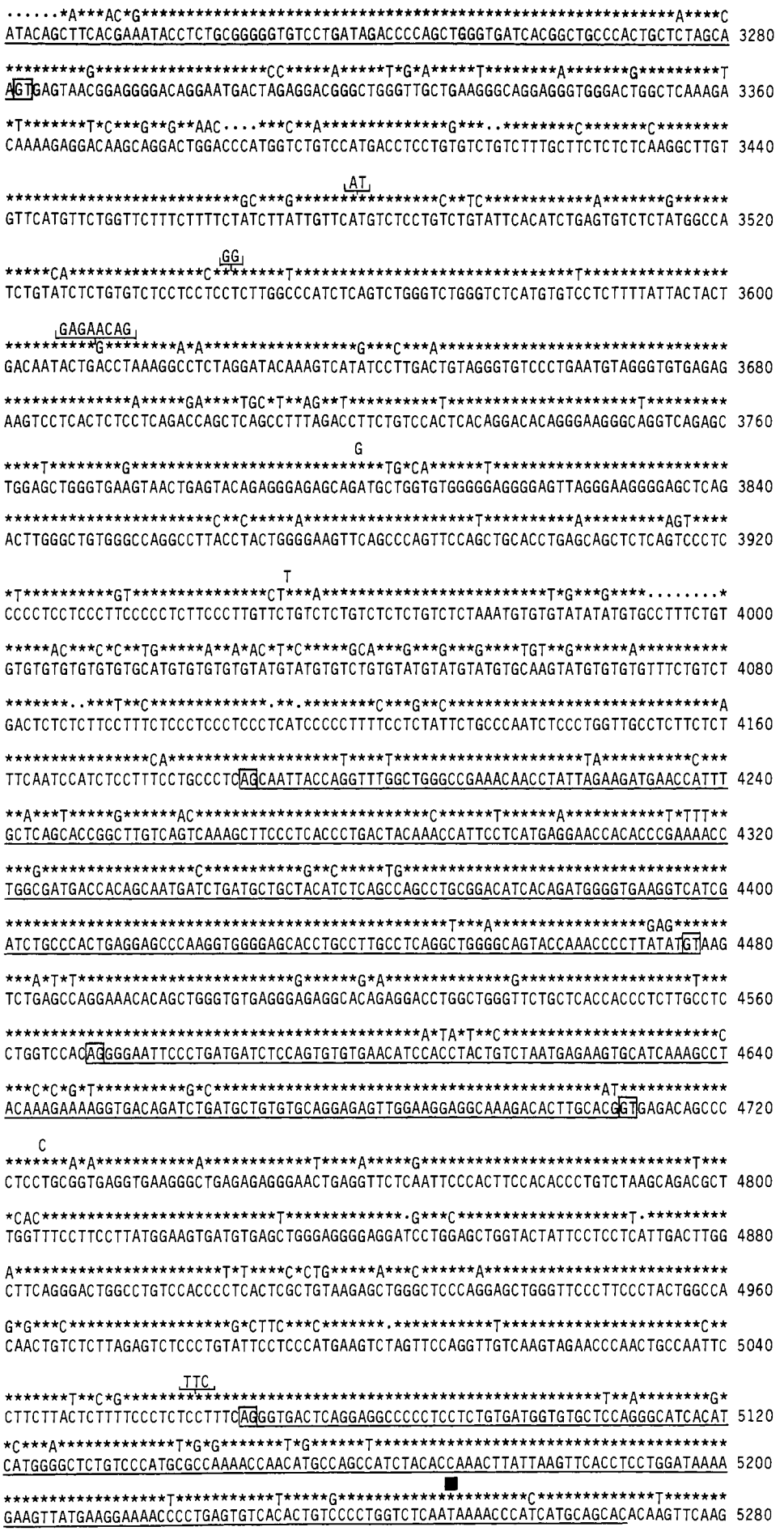
Expression of RSKG-7 and RSKG-3. The sequences of RSKG-7 and RSKG-3 transcripts share 85% identity. Most of the heterogenity between these two genes is attributed to several regions of hypervariable sequences along the transcripts. Oligonucleotide probes complementary to two of those hypervariable regions were synthesized and end labeled with ^{32}P. One is RSKG-7-specific 21-mer (5′ TCCCATA-TAAGGGGTTTGGTA 3′) complementary to the junctional region of the third and the fourth exon; the second is RSKG-3-specific 21-mer (5′ GCAGAGGTATCTGTTGATGAC 3′) complementary to a region of the second exon where six nucleotides were deleted. Northern blot analyses, using these specific probes, demonstrate that RSKG-7 is expressed in both submandibular gland and kidney at comparable levels (Figure 5), but it is not readily detectable in the liver, heart, lung, pancreas, spleen, brain, pituitary gland, thymus, adrenal gland, and testis. RSKG-3 mRNA is also only detectable in the submandibular gland and kidney among the 12 tissues analyzed, but its expression level in those two tissues was much lower than that of RSKG-7. Dot blot studies show that the amount of the RSKG-3 mRNA in the submandibular gland is an order of magnitude higher than that in the kidney (data not shown). Using the tissue kallikrein specific oligonucleotide probe for a comparative study, we found that tissue kallikrein was expressed as previously reported in the pancreas, salivary gland, and kidney (MacDonald & Ashley, 1985). Additionally, we demonstrate here the expression of tissue kallikrein in the spleen and prostate (Figure 5).

Effects of Testosterone and Thyroxine on the RSKG-7 mRNA Level in Castrated Male Rats. Dot blot analyses using a specific oligonucleotide probe show that castration results in a decrease of the submandibular gland RSKG-7 mRNA level, which can be restored to normal by testosterone or thyroxine treatment (Figure 6). Castration, on the other hand, has no effect on the kidney RSKG-7 mRNA level, nor does testosterone or thyroxine treatment (Figure 6). Chao and Margolius (1983) have reported that testosterone and thyroxine modulate both kallikrein-like activity and quantity in the castrated male and normal female rat submandibular gland but not in the kidney. Our study on the hormonal control of the expression of RSKG-7 in male rats is consistent with these results.

Discussion

We have sequenced and characterized two rat renal kallikrein genes. The coding sequences among RSKG-7, RSKG-3, and tissue kallikrein share more than 85% identity. The noncoding and the 5′ and 3′ flanking sequences of these genes


RSKG-3
RSKG-7



```
**************GA*********-------------------------------------------------------------
TTCTGATTCTCTCTCACCTTCTGGAGTCTGGAGCTGGGGTCTGCCCAGCTAAGGTCAGGACTTCCAAAGCAGATTGGCAA 5360
-------------------------------------------------------------------------------
TGGATAGGGGTCCATGTATGGTTGGATAAAGGGATTGGCTTGAGACAAAGGCAGACAAAAGGCTCAGACCATGTCAGGGA 5440
-------------------------------------------------------------------------------
TGGCCCAGGGTTCTCCTCTGGCAAGTAGCATGTTCAGATCACAAAGTTTAGTCAAGTTAGACACAGTAGAAAGATCAACT 5520
-------------------------------------------------------------------------------
CCAGCTAGGCCGGACCCTACACCAAAAGCAGAAGAGATGACAAGCAGTCAGGACACTGAGAATTTAGACTGCCGACCTGC 5600
-------------------------------------------------------------------------------
TTTTATCGAGTCGGGGAAGCCAGAAAGATGGCGGGATGTTATATTAATTAAGCCTTTACTTCTTGGGGAGTTGTCTGTCC 5680
-------------------------------------------------------------------------------
ACAGAAGGGTAGGTACTGCCCCCCAAGGCTAGGCTGCAGAGTATCTCCTAGAGCCCCTGGAACCCCTGGAGCCCCTGGAG 5760
-------------------------------------------------------------------------------
CCCCTGGAGCCCCTGGAACCCCTGGAGCCCCTGGAGCCCCTGGAGGGTATATCTCTGGTAGGGCTGGACCAAAGATGAAG 5840
-------------------------------------------------------------------------------
GGTGACTTAGCGGCATGCCCTGTGGCAGAGCATCGCCCGGCTTGGGAAAAGGGCCACATTGCTCCCACTGCCTTGGACTA 5920
-------------------------------------------------------------------------------
CAAACCAGTGTGTGCCCAGAGTGATATAGCCGCAGCTGAGAGTATTGGAGGACTTTCAGAGGAGGGTCCCTGCCAAGACT 6000
-------------------------------------------------------------------------------
CAGGCAGTAAGGAGTGGGCCCTAGAGTGGGGACTGGGGCTTGCCACCTGTGTCCCCAGGGTGCTACCTGCTCTCAGGGGA 6080
-------------------------------------------------------------------------------
AGTATGGGACCTCAAGGAAGATCTGATGGAAACTGAGCCACAGGGACCGAGGTAGAGGCATCAGGCTGCGGGGACTCATC 6160
-------------------------------------------------------------------------------
AGTGTCCTGAGAAGTTGCGTCGTCCCTCTGTTGGGGAGGCTCGAGGTATTTCACAAAATTGCTTGCAGTAGACAGCTTCT 6240
----------------
GGGCGGGCTCCAAGCTT
```

FIGURE 2: Comparison of nucleotide sequences of RSKG-7 and RSKG-3. Only the sense strand is shown. The five exons of the gene are underlined. The variant CATAT box and TTTAAA box of RSKG-7 are double-underlined. The transcriptional initiation site of RSKG-7 is marked by a triangle. The RSKG-3 exon sequence upstream of a *Eco*RI site (denoted by a dagger) in the second exon was determined by mRNA dideoxy sequencing. The available RSKG-3 sequence is aligned with RSKG-7 for highest percentage of identity. The unsequenced region of RSKG-3 is indicated by dashes. The nucleotides of RSKG-3 identical with RSKG-7 are denoted by asterisks. The insertion sequences of RSKG-3 are looped out, and the deletions are shown by dots. The filled square indicates the polyadenylation signal of both genes. The splicing sites around each exon are boxed.

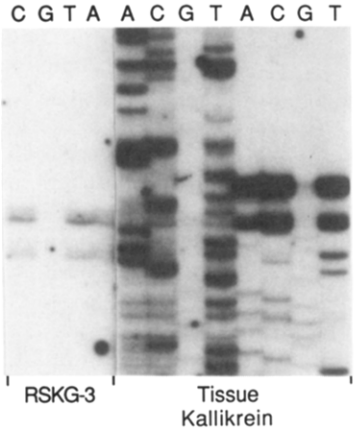


FIGURE 3: Mapping of the transcriptional initiation site of RSKG-3 and tissue kallikrein. Messenger RNA dideoxy sequencing using a primer (5′ TCCAATCCGTCAGGTGTGATG 3′) specifically complementary to the bases 513–533 of tissue kallikrein is used as a length marker (center four lanes). RSKG-3 mRNA sequencing was primed by a 21-mer (5′ GCAGAGGTATCTGTTGATGAC 3′) that is complementary to a region of the second exon where six nucleotides are deleted from RSKG-3 compared to the tissue kallikrein gene; tissue kallikrein mRNA sequencing reaction was primed by another oligonucleotide (5′ ACATAGGTATTCGCCGAAGTA 3′) complementary to the corresponding region.

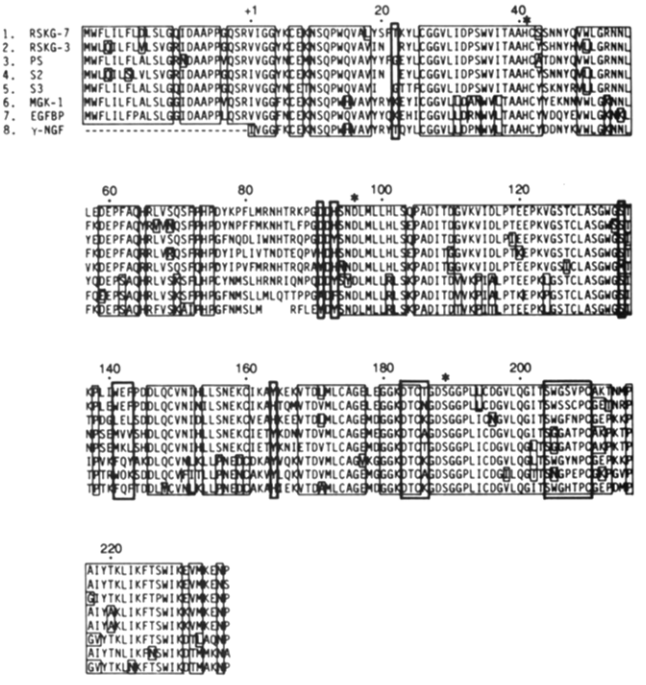


FIGURE 4: Amino acid sequence comparison of RSKG-7, RSKG-3, and the kallikrein-like proteins. The predicted amino acid sequences of RSKG-7 and RSKG-3 are aligned with the amino acid sequences of PS (rat pancreatic kallikrein), S2 (rat tonin), and S3 (Ashley & MacDonald, 1985), MGK-1 (Mason et al., 1983), EGF-BP (epidermal growth factor-binding protein) (Lundgren et al., 1984), and γ-NGF (γ subunit of nerve growth factor) (Thomas et al., 1981). The numbering starts at the amino terminus of the predicted active form of rat kallikrein. The boxed areas delineate regions of the sequences that share a high percentage of identity. The residues believed to line the substrate-binding pocket are heavily boxed. The asterisks denote the positions of the amino acids forming the catalytic triad.

are also substantially homologous, suggesting that the kallikrein gene family has evolved from relatively recent gene duplication events. The coding sequence of RSKG-7 is approximately one-fifth of the total length of the gene. It is interrupted by four intervening sequences, 1847, 907, 96, and 370 bp, respectively for the first, second, third, and fourth intron, with the exon–intron arrangement closely resembling that of the mouse kallikrein gene. RSKG-3 has an overall gene organization identical with that of RSKG-7. Interestingly, 40–50 bp of poly(dG-dT):poly(dA-dC) sequence are found near the end of the second intron of both RSKG-7 and RSKG-3 (Figure 2). A perfect stretch of 31 bp of poly(dT-dC):poly(dG-dA) is located in the 3′ end of the first intron of RSKG-7. Eucaryotic genomes contain numerous runs of alternating deoxyguanosine and thymidine residues (GT elements), 10–50 bp in length (Hamada & Kakunaga, 1982; Hamada et al., 1982), whereas TC elements only appear in the genomes of some species (Hamada et al., 1984a). Repeated sequences containing d(GT/AC) are known to form

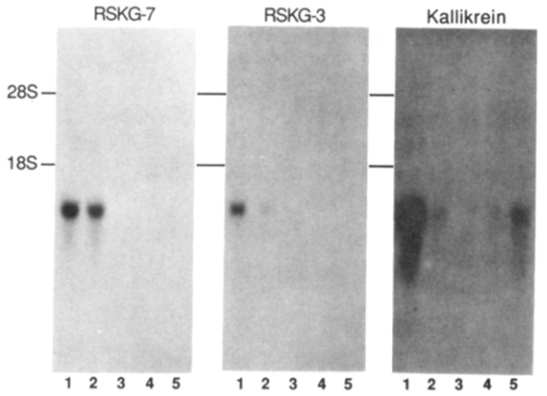


FIGURE 5: Tissue-specific expression of RSKG-7 and RSKG-3: (Lane 1) Salivary gland; (lane 2) kidney; (lane 3) prostate; (lane 4) spleen; (lane 5) pancreas. One hundred micrograms total RNA of each tissue was electrophoresed through a formaldehyde–agarose (1.5%) gel, transferred to Zeta-probe membrane, and hybridized to a ^{32}P-labeled synthetic oligonucleotide probe specific to RSKG-7, RSKG-3, or tissue kallikrein.

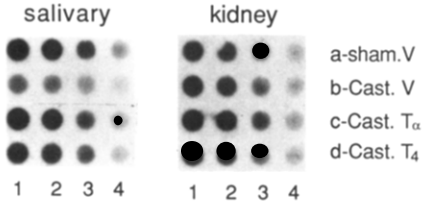


FIGURE 6: Testosterone and thyroxine induction of RSKG-7 expression. Four groups of male rats were used in this study: (a) sham-operated rats; (b) castrated rats; (c) castrated rats treated with testosterone; (d) castrated rats treated with thyroxine. Dot blots of total RNA: (1) 25 μg, (2) 12.5 μg, (3) 6.2 μg, or (4) 3.1 μg of salivary gland or kidney from the male rats with different treatments were hybridized to an RSKG-7-specific oligonucleotide probe and autoradiographed for 3 days at −70 °C.

a left-handed Z-DNA conformation in vitro (Singleton et al., 1982) and in vivo (Nordheim et al., 1981). GT elements also have been implicated in DNA recombination (Stringer, 1985) and weak enhancer activity (Hamada et al., 1984b). Both GT and TC elements are found in the mouse IgG2a heavy chain constant region gene (Cγ2a) and have been implicated in unequal sister chromatid exchange via homologous recombination (Weinerb et al., 1988). The rat tonin gene, which has been sequenced, also contains the GT and TC elements at the corresponding regions (Shai et al., unpublished results). GT and TC elements may play a role in the intergenic recombination and gene conversion events by which the high degree of sequence homology of the kallikrein-related genes is maintained.

Dideoxy sequencing method was used in order to achieve both sequencing and 5′ end mapping of RSKG-3 mRNA at the same time. Figure 3 shows there are a total of five strong sequencing terminations clustered in two regions toward the 5′ end of both RSKG-3 and tissue kallikrein transcripts. The top band of each of the two clusters corresponds to those detected by the primer extension assay, and they were previously described as two putative transcriptional initiation sites of the mouse kallikrein gene (Mason et al., 1983). The five strong terminations, with the exception of the mature runoff, are unlikely to be additional transcriptional initiation sites of RSKG-3 and the tissue kallikrein. The appearance of the missing bands and the nonspecific terminations are likely due to the fact that transcription of the 5′ end of the kallikrein-related mRNAs by the reverse transcriptase are error prone, which may explain why the first putative transcription initiation site was not reached by the kallikrein cDNA cloning and sequencing study of Ashley and MacDonald (1985). Therefore, we suggest there is only one transcription initiation site for RSKG-3, tissue kallikrein, and RSKG-7. Consequently, the 5′ end untranslated region of RSKG-7, RSKG-3, and tissue kallikrein transcripts is 42 nucleotides instead of the putative 36 or 42 bases. The high degree of sequence and structural homology among the kallikrein-related genes implies that a single transcriptional initiation site could be a general phenomenon of the kallikrein gene family.

We have identified a deletion of six nucleotides in a hypervariable region of the second exon of RSKG-3 (Figure 2). This deletion involves sequences coding for amino acids that participate in substrate binding. Ashley and MacDonald (1985) also found two kallikrein-related cDNAs, S2 and S3, with a six-nucleotide deletion at the corresponding region. These two cDNAs share 74–80% sequence identity with RSKG-3, suggesting that these three genes belong to a main subgroup of the kallikrein gene family.

Both RSKG-7 and RSKG-3 encode a precursor of kallikrein-like proteins destined for secretion, and there is only one potential glycosylation site (ASN-X-SER/THR) in each of the proteins (Figure 4). Amino acid residues histidine-41, aspartate-96, and serine-189 of the serine protease charge relay system and their contiguous sequences are conserved in both RSKG-7 and RSKG-3 (Figure 4). The aspartate residues at position 183 encoded by both genes are characteristic of proteases with a trypsin-like cleavage preference; however, the insertion of a proline at position 209 increases the size of the substrate-binding pocket of the encoded proteases over that of trypsin (Bode et al., 1983). Amino acid residues at positions 206 and 217 are located at the mouth of the substrate-binding pocket in the porcine kallikrein to accommodate bulky amino acid side chains (MacDonald & Ashley, 1985). The presence of small amino acids (glycine, alanine, or serine) at the corresponding region in RSKG-7 and RSKG-3 also ensures an open substrate-binding pocket in RSKG-7 and RSKG-3. Chen and Bode (1983) have proposed that a principal determinant in the specific cleavage of the kininogen precursor by kallikrein is the presence of a hydrophobic sandwich formed by tyrosine-93 and tryptophan-205 that facilitates the binding of bulky, hydrophobic residues in the substrate adjacent to the cleavage site. Both RSKG-7 and RSKG-3 encode histidine and tryptophan in the corresponding positions, indicating that they may also have kininogenase activity. However, despite the similarity of the framework of the enzymes, the key amino acids that have been implicated in contacting the substrates are significantly different among RSKG-7, RSKG-3, tissue kallikrein, tonin, EGF-BP, and γ-NGF (Figure 4), which strongly suggests that RSKG-7 and RSKG-3-encoded serine proteases may have unique substrate specificities.

The tissue-specific expression of kallikrein-related genes was analyzed by Northern hybridization using specific synthetic oligonucleotide probes. Both RSKG-7 and RSKG-3 transcripts are only detectable in the submandibular gland and kidney among the 12 tissues studied. RSKG-7 is expressed at the same level in both submandibular gland and kidney. On the other hand, RSKG-3 is expressed at an order of magnitude less in the kidney than in the submandibular gland. Quantitation of the gene expression in kidney by dot blot demonstrates that RSKG-7 is expressed 5-fold higher than tissue kallikrein and 20-fold higher than RSKG-3. The expression spectrum of RSKG-7 and RSKG-3 is much narrower than that of tissue kallikrein; in a parallel study, tissue kallikrein is found to be expressed in the submandibular gland, pancreas, and kidney in decreasing order of magnitude, which is con-

sistent with previously published results (MacDonald & Ashley, 1985). Additionally, we have found that tissue kallikrein is also expressed in spleen and prostate gland. RSKG-7 and RSKG-3 encode two serine proteases that are different from two other kallikrein-related proteins, tissue kallikrein and tonin (Woodley-Miller et al., 1987), expressed in the kidney. Whether either of these two genes code for arginine esterase A (Chao, 1983) is still under investigation. Our results indicate that there are at least four or five kallikrein-like proteins, RSKG-7, RSKG-3, tissue kallikrein, tonin, and arginine esterase A, being produced in the rat kidney.

Our previous experiments with hormonal manipulation of rats indicate that tissue kallikrein mRNA levels in the submandibular gland are responsive to hormonal regulation (Gerald et al., 1986b). In the present study, testosterone and thyroxine were shown to be effective in restoring the submandibular RSKG-7 mRNA level of castrated male rats to normal. On the other hand, neither hormonal treatment nor castration affects the RSKG-7 mRNA level in the kidney. Whether the changes in RSKG-7 mRNA levels are direct or indirect effects of the hormonal manipulations remains to be investigated.

Kallikrein-related proteins have been implicated in the processing of vasoactive peptides and growth factors (Schachter, 1980). Tissue kallikrein cleaves kininogen to release kinin, which has been implicated in local blood flow and water or ion transport in kidney (Levinsky, 1979). The "kidney-specific" expression of RSKG-7 and RSKG-3 implies that they are potential candidates to participate in the homeostasis of the kidney. The characterization of the substrate specificity of these two new enzymes awaits their purification from rat kidney or from cell lines that are transformed to express RSKG-7 and RSKG-3 gene products by recombinant DNA techniques.

REFERENCES

Ashley, P. L., & MacDonald, R. J. (1985) *Biochemistry 24*, 4512–4520.

Berger, E. A., & Shooter, E. M. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 3647–3651.

Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt-Kastner, G., & Bartunik, H. (1983) *J. Mol. Biol. 164*, 237–282.

Boucher, R., Asselin, J., & Genest, J. (1974) *Circ. Res., Suppl. 34*(1), 203–209.

Chao, J. (1983) *J. Biol. Chem. 258*, 4434–4439.

Chao, J., & Margolius, H. S. (1983) *Endocrinology (Baltimore) 113*, 2221–2225.

Chao, J., Woodley, C., Chao, L., & Margolius, H. S. (1983) *J. Biol. Chem. 258*, 15173–15178.

Chao, J., Chao, L., & Margolius, H. S. (1984a) *Biochim. Biophys. Acta 801*, 244–249.

Chao, J., Shimamoto, K., & Margolius, H. S. (1984b) *Hoppe-Seyler's Z. Physiol. Chem. 365*, 1137–1141.

Chao, J., Chao, L., Swain, C., Tsai, J., & Margolius, H. S. (1987) *Endocrinology (Baltimore) 120*, 475–482.

Chen, Z., & Bode, W. (1983) *J. Mol. Biol. 164*, 283–311.

Davis, L. G., Dibner, M. D., & Battey, J. F. (1986) *Methods Mol. Biol.*, 130–146.

Evans, B. A., & Richards, R. I. (1985) *EMBO J. 4*, 133–138.

Frey, P., Forand, R., Maciag, T., & Shooter, E. M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6294–6298.

Gerald, W. L., Chao, J., & Chao, L. (1986a) *Biochim. Biophys. Acta 866*, 1–14.

Gerald, W. L., Chao, J., & Chao, L. (1986b) *Biochim.* Biophys. *Acta 867*, 16–23.

Hamada, H., & Kakunaga, T. (1982) *Nature (London) 298*, 396–398.

Hamada, H., Petrino, M. G., & Kakunaga, T. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6465–6469.

Hamada, H., Petrino, M. G., & Kakunaga, T. (1984a) *Mol. Cell. Biol. 4*, 2622–2630.

Hamada, H., Petrino, M. G., Kakunaga, T., Seidman, M., & Stollar, B. D. (1984b) *Mol. Cell. Biol. 4*, 2610–2621.

Hamlyn, P. H., Brownlee, G. G., Cheng, C. C., Gait, M. J., & Milstein, C. (1978) *Cell (Cambridge, Mass.) 15*, 1067–1075.

Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L., & Steitz, J. A. (1980) *Nature (London) 283*, 220–224.

Levinsky, N. G. (1979) *Circ. Res. 44*, 441–451.

Lundgren, S., Ronne, H., Rask, L., & Peterson, P. A. (1984) *J. Biol. Chem. 259*, 7780–7784.

MacDonald, R. J., & Ashley, P. L. (1985) *Biochemistry 24*, 4520–4527.

Margolius, H. S. (1984) *Annu. Rev. Physiol. 46*, 309–326.

Margolius, H. S., Morowitz, D., Pisano, J. J., & Keiser, H. R. (1974) *Circ. Res. 35*, 820–825.

Mason, A. J., Evans, B. A., Cox, D. R., Shine, J., & Richards, R. I. (1983) *Nature (London) 303*, 300–307.

Neurath, H., & Walsh, K. A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 3825–3832.

Nordheim, A., Pardue, M. L., Lafer, E. M., Moller, A., Stollar, B. D., & Rich, A. (1981) *Nature (London) 294*, 417–422.

Sanger, F., Nicklen, S., & Coulson, A. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Schachter, M. (1980) *Pharmacol. Rev. 31*, 1–17.

Schreir, P. H., & Cortese, R. (1979) *J. Mol. Biol. 129*, 169–172.

Singleton, C. K., Klysik, J., Stirdivant, S. M., & Wells, R. D. (1982) *Nature (London) 299*, 312–316.

Sood, A. K., Pereira, D., & Weissman, S. M. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 616–620.

Stringer, J. R. (1985) *Mol. Cell. Biol. 5*, 1247–1259.

Swift, G. H., Dagorn, J.-C., Ashley, P. L., Cummings, S. W., & MacDonald, R. J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7263–7276.

Thomas, K. A., Baglan, N. C., & Bradshow, R. A. (1981) *J. Biol. Chem. 256*, 9156–9166.

Tschesche, H., Mair, G., Godec, G., Fielder, F., Ehret, W., Hirschauer, C., Lemon, M., & Fritz, H. (1979) *Adv. Exp. Med. Biol. 120*, 245–260.

Weinerb, A., Katzenberg, D. R., Gilmore, G. L., & Birshtein, B. K. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 529–533.

White, B. A., & Bancroft, F. C. (1982) *J. Biol. Chem. 257*, 8569–8572.

Woodley-Miller, C., Chao, J., & Chao, L. (1987) *Biochem. J. 248*, 477–481.